Attorney Docket: 1278-20

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): Alexander A. Kist

Serial No: 10/805,944

Filed: May 24, 2007

Dated: September 26, 2008

For: ESTIMATING BANDWIDTH

Commissioner of Patent
P.O. Box 1450
Alexandria, VA 22313-1450

LETTER

Enclosed is a certified copy of Australian application no. 2006-902805 filed May 24, 2006 and from which priority is claimed under 35 U.S.C. § 119.

Respectfully submitted,

George M. Kaplan
Registration No. 28,375
Attorney for Applicant

DILWORTH & BARRESE, LLP
333 Earle Ovington Blvd.
Uniondale, NY 11553
(516) 288-8484

CERTIFICATE OF MAILING UNDER 37 C.F.R. §1.8(a)

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail, postpaid in an envelope, addressed to the: Mail Stop Missing Parts, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on September 26, 2008

September 26, 2008
Dated

George M. Kaplan

Australian Government

Patent Office
Canberra

I, JANENE BRYDE, MANAGER EXAMINATION SUPPORT AND SALES hereby certify that annexed is a true copy of the Provisional specification in connection with Application No. 2006902805 for a patent by AUSTRALIAN TELECOMMUNICATIONS COOPERATIVE RESEARCH CENTRE as filed on 24 May 2006.

COMMONWEALTH OF AUSTRALIA
AUSTRALIA
PATENT OFFICE

WITNESS my hand this
First day of June 2007

J. K. Bryde

JANENE BRYDE
MANAGER EXAMINATION SUPPORT
AND SALES

P/00/009
Regulation 3.2

AUSTRALIA

Patents Act 1990

PROVISIONAL SPECIFICATION

Invention Title: **Estimating bandwidth**

The invention is described in the following statement:

ESTIMATING BANDWIDTH

Field of the Invention

The present invention relates generally to a method for estimating bandwidth available on networked interfaces. In a preferred form the invention is applied to estimating available bandwidth on networked interfaces and using those estimations to optimise routing decisions for data packets being sent through those networked interfaces.

Background

The Internet includes a large number of private networks owned by companies, organisations or individuals. Each of these private networks have at least one interface connecting the Private network to the Internet, with many private networks having more than one such interface.

Where only one interface from the private network to the Internet exists all outgoing data packets being sent from a host on the private network to a host on the Internet must pass through that interface. Where multiple interfaces exist, however, the private network must decide on which interface the outgoing data packet should be sent.

Where multiple interfaces exist an approach to routing outgoing traffic is to have one default interface through which all traffic is sent, and alternative interfaces which are only used in the event that the default interface has no more capacity to send additional packets (i.e. an overflow scenario) or in the event that the default interface fails (i.e. a failover scenario).

Such an arrangement often fails to optimise the use of multiple external interfaces. Optimisation, for example, may be in terms of connection quality, even distribution of outgoing traffic, or optimisation of costs associated with different interfaces.

In order to make more strategic routing strategies one approach has been to estimate the bandwidth available on external interfaces and make routing decisions accordingly. One way such estimations have been achieved is by analysing the amount of data being sent out in a particular data flow. By looking at

2006902805 24 May 2006

data packets travelling from the same source to the same destination a prediction of the continued size of the flow is made, and the available bandwidth of the external interface on which that flow is travelling through is updated accordingly. In this way a prediction may be made as to the amount of traffic that is being and will, in the near future, be sent through a particular interface, and with this information routing decisions may be made.

While performing such predictions provides for a more accurate estimation of used and available bandwidth than a mere consideration of data packets being sent without forecasting future traffic, it may be advantageous to have an alternative and preferably more accurate method for such estimations.

Any reference in this specification to the prior art does not constitute an admission that such prior art was well known or forms part of the common general knowledge in any jurisdiction.

Summary of the invention

The present invention is based on an insight that prior art systems have ignored terminating data flow components when implementing or developing routing schemes, and policies in the past and that by addressing this oversight by including a consideration of terminating flow components in bandwidth estimations an improved allocation of data transmission resources may result.

Accordingly in a first aspect the present invention provides a method of estimating used or available bandwidth for a network interface, the network interface providing data connectivity between a first network node and a second network node, the method including estimating bandwidth use on the basis of data transmitted, in forward and reverse directions between the first network node and the second network node.

Data transmitted in the forward direction can be nominally defined as data emanating from the first node and terminating at the second node, and data transmitted in the reverse transmission direction can be defined as data terminating at the first node that emanated from the second node.

In a preferred form, the method includes making separate estimates of bandwidth use for data being transferred in the forward direction and data being transferred in the reverse direction.

Preferably the forward bandwidth estimate is calculated on the basis of data packets being sent from the first network node to the second network node. Similarly a reverse bandwidth estimate is calculated on the basis of data packets being sent from the second network node to the first network node.

In a second aspect the present invention provides a method of estimating the bandwidth of a data flow between a first network node and a second network node of a network, the data flow including a forward flow component emanating from the first network node and terminating at the second network node and a reverse flow component emanating from the second network node and terminating at the first network node, the method including:

estimating the bandwidth of the forward flow component on the basis of one or more data packets belonging to the forward flow;

estimating the bandwidth of the reverse flow on the basis of one or more data packets belonging to the reverse flow.

The method can include estimating the bandwidth of either the forward or reverse flows on the basis of one or more of the following flow component parameters:

a size of one or more data packets;

a transmission frequency of data packets belonging to the flow component;

the total amount of data transmitted that belongs to the flow component;

the amount of data transmitted in a predetermined time period that belongs to the flow component;

a total number of data packets belonging to the flow component that have been transmitted;

a number of data packets transmitted that belong to the flow component;

an average size of a data packet belonging to the flow component.

The method can include determining an aggregate estimate of the bandwidth of the flow by combining the estimates for the forward flow component and reverse flow component.

In broad concept a further aspect of the present invention provides a method of assigning a bidirectional dataflow to one of a plurality of network interfaces on the basis of estimated forward and reverse bandwidth requirements of the flow.

In a further aspect the present invention provides a method of assigning to a data flow a transmission interface, from a plurality of such interfaces, through which one or more data packets belonging to the dataflow will be routed, the method including:

identifying at least one data packet belonging to the flow;

estimating the flow's forward and reverse bandwidth; and

assigning the dataflow for transmission on one of the transmission interfaces on the basis of the flow's estimated the flow's forward and reverse bandwidth.

The assignment may be performed in accordance with a predetermined optimisation algorithm.

The optimisation algorithm may specify a default transmission interface amongst the plurality of interfaces. In the event that the flow's estimated forward and reverse bandwidth is incompatible with assignment to the default interface the optimisation algorithm can assign the dataflow to be transmitted on an alternative transmission interface.

The optimisation algorithm can be configured to perform assignment of a flow to an interface of the plurality of interfaces to optimise at least one of cost, quality of service, speed of transmission.

The step of estimating the flow's forward and reverse bandwidth can be performed in accordance with an embodiment of the first aspect of the present invention.

Another aspect of the invention provides a method of packet forwarding a data packet received on a first interface of a network to one of a plurality of second interfaces that connect the network to at least one second network, wherein each data packet received at the first interface belongs to a data flow between a first network node in data communication with the first interface and a second network node in data communication with at least one of the second interfaces, the method including:

determining whether a data packet received at a first interface belongs to a known flow; and in the event that the received data packet belongs to an unknown flow,

making an initial estimate of the flow's forward and reverse bandwidth; and

forwarding the data packet on one of the second interfaces on the basis of the flow's estimated forward and reverse bandwidth.

A further aspect of the invention provides a method in a device, capable of monitoring data being transmitted between a first network and second network via a plurality of communication interfaces; the method including;

detecting the arrival of a data packet at an interface;

determining a dataflow to which the packet belongs;

estimating a forward and reverse bandwidth of the dataflow to which the packet belongs; and

assigning a proportion of an available forward and reverse bandwidth on one of the interfaces for the transmission of the dataflow on the basis of the estimate.

In the event that the packet emanates from the second network, and the data packet is identified as belonging to a new dataflow, the method can additionally include assigning the interface on which the data packet was detected as the interface on which to transmit all future packets belonging to the dataflow.

A still further aspect of the invention provides a network device configured to implement any one of the methods described herein.

In another aspect the present invention provides a network device configured to estimate used or available bandwidth of a network interface providing data connectivity between a first network node and a second network node, the network device including means to estimate bandwidth use on the basis of data transmitted, in forward and reverse directions, between the first network node and the second network node.

Preferably the means to estimate bandwidth use generates an estimate of bandwidth use in a forward direction and an estimate of bandwidth use in a reverse direction.

In one form the network device forms is mounted in the data transmission path between the first network node and a second network node. Alternatively the network device does not form part of the data transmission path between the first network node and a second network node. In this case the network device may receive data representing the flow of data between the first network node and a second network node from which estimates can be made.

A network device for determining packet routing information for a data packet belonging to a data flow being sent from an internal network to an external network, the external network connected to the internal network through at least one of a plurality of interfaces, the device including: means for receiving data characterising the data packet; means for determining from at least part of the received information whether the data packet belongs to a known data flow; means for estimating the forward and reverse bandwidth of the data flow to which the data packet belongs on the basis of at least part of the received information; and means for determining which of the plurality of interfaces the data packet should be routed.

The network device can include means for identifying a new data flow. In this case, the device can further include means for assigning an interface to the new data flow on which packets belonging to the flow will be routed.

2006902805 24 May 2006

In yet another aspect the invention relates to a router including a network device for determining packet routing information according to an embodiment of the previous aspect of the invention.

In a further aspect the present invention provides a system including a plurality of network nodes coupled to a data communications network, the network further including at least one port for communicating with a plurality of interfaces through which data communications with at least one external network are enabled, the system further including a network device configured to estimate used or available bandwidth of at least one of the interfaces on the basis of data transmitted, in forward and reverse directions, between a network node of the system and a network node of an external network.

In another aspect the present invention provides a system including a plurality of network nodes coupled to a data communications network, the network further including at least one port for communicating with a plurality of interfaces through which data communications with at least one external network are enabled, the system further including a network device for determining packet routing information for a data packet belonging to a data flow being sent from an a network to the external network, the device including: means for receiving data characterising the data packet; means for determining from at least part of the received information whether the data packet belongs to a known data flow; means for estimating the forward and reverse bandwidth of the data flow to which the data packet belongs on the basis of at least part of the received information; and means for determining which of the plurality of interfaces the data packet should be routed.

In embodiments of the previous embodiments the system can include a router and/or packet forwarder to forward data on a selected one of the interfaces.

2006902805 24 May 2006

Brief description of the drawings

An illustrative embodiment of the present invention will now be described by way of non-limiting example only with reference to the accompanying drawings. In the drawings:

Figure 1 provides a schematic illustration of a network, including a private network operating in accordance with an exemplary embodiment of the present invention and linked via three interfaces to the internet;

Figure 2 depicts the data flow fields stored in a flow tracker operating according to an embodiment of the present invention; and

Figure 3 depicts a flowchart illustrating an process for routing data packets that uses an exemplary embodiment of at least one aspect of the present invention and which may be implemented in a system such as that of figure 1.

Detailed description

Figure 1 illustrates a private network 100 connected to the Internet 102 through a number of different Internet Service Provider interfaces 104, 106, 108. The private network 100 includes a number of nodes, for example node x1 110, node x2 112 and node x3 114, resident on the private network. In addition to traditional network hardware and software components (not shown) the private network 100 also includes a Routing Management Application (RMA) module 116. The Internet includes a number of nodes, for example node Y1 118, node Y2 120 and node Y3 122. Each node on the private network 100 can connect to the Internet 102, and is therefore connectable to each node on the Internet 102, though any one of the interfaces 104, 106 and 108.

All data packets being sent from a node 110, 112 or 114 inside the private network 100 to a destination host 116, 118 or 120 outside the private network (on the Internet 102) – i.e. emanating packets, must be routed through one of the interfaces 104, 106 or 108. For the purposes of the preferred embodiment it will be presumed that any destination host 118, 120 or 122 on the Internet 102 may be reached through any of the interfaces 104, 106 or 108 – i.e. the decision of which

interface 104, 106 or 108 to route an emanating packet through is independent of the destination of that packet. Any data packets being sent from a host 118, 120, 122 outside the private network 100 to a host on the private network 100 (i.e. terminating packets) must enter the private network 100 through one of the interfaces 104, 106, 108.

Because the private network 100 possesses links to three such interfaces it has the ability to manage how packets are routed through the interfaces. The decision on how routing of particular data packets should be performed may be dependent on many factors such as the cost of using each of the interfaces, the available bandwidth on each of the interfaces, the quality of service or data transfer speed provided by each of the interfaces and other factors which may influence a routing decision that will be appreciated in those skilled in the art.

In order to manage the routing of data packets, and more particularly data flows the RMA module 116 is provided.

The RMA module 116 intercepts, analyses and routes all packets emanating from the private network 100 to one of the available interfaces 104, 106 or 108. Similarly, all terminating packets entering the private network 100 through one of the interfaces 104, 106 or 108 are intercepted by the RMA 116 for analysis prior to being forwarded to the final destination of the packet (one of the nodes 110, 112 or 114). The RMA 116 functions to implement flow tracking, bandwidth management, flow based routing strategies, failover, and capacity discovery, each of which will be described in detail below.

For the purpose of the discussion of the preferred embodiment data packets will be deemed to be either emanating or terminating (being sent from, or to, the private network 100). Emanating data packets are sent form a source host (node) 110, 112 or 114 on the private network 100 to a destination host (node) 118, 120 or 122 on the Internet 102. In the preferred embodiment, emanating packets are sent from a host e.g. node, 110. and intercepted by the RMA 116 before being forwarded to an interface 104, 106 or 108 for routing to a destination, e.g. host 118.

Terminating data packets are sent form a source host 118, 120 or 122 on the Internet 102 to a destination host 110, 112 or 114 on the private network 100. In the preferred embodiment terminating packets are received by one of the interfaces 104, 106 or 108 and are passed to the RMA 116 before being forward to a destination host, e.g. node 110, on the private network 100.

In the examples described below data packets being transmitted between two network nodes, node X1 110 and node Y1 118, will be discussed. In this case:

- node 110 is deemed to have an address of 'x' and node 118 is deemed to have an address of 'y' emanating packets will therefore have a source address x, and a destination address y; and
- terminating packets will have a source address y and a destination address x.

The full source and destination information of a data packet may additionally include information such as an IP address, a port address and/or a protocol identifier.

Data Flows

The RMA relies on the concept of data flows to analyse network traffic. Traditionally, data flows are considered to be the aggregate of data packets being sent from the same source to the same destination – i.e. all emanating packets being sent from source x to destination y. Although this traditional approach is useful, it only provides a part picture of data communications in a packet switched environment.

In the majority of cases, packet switched communication involves information being transmitted in two directions. Information is sent from a source x to a destination y (emanating data) and in response to that information the destination y sends information back to the original source x (terminating data). This response information may simply be acknowledgment data. Alternatively, the emanating data may be a request for data (such as a file, a web page, streaming audio/video), in which case a response including the requested data will be sent

back to the source of the emanating data from the destination of the emanating data.

To account for this two way flow of information, data flows in the preferred embodiment of the invention include an emanating flow component and a terminating flow component both of which are considered in making bandwidth estimations and flow routing decisions. The emanating flow component includes all data packets being sent from source X1 110 to destination Y1 118, and the terminating flow component includes all data packets being sent back to source X1 110 from source Y1 118. The emanating flow, for example, may comprise data packets sent from source X1 110 requesting information from destination Y1 118. In this case, the terminating flow is the data packets being sent from Y1 118 back to X1 110 in response to the initial request from X1 110.

In order to identify different data flows and associate a particular data packet with a particular data flow, when the RMA 116 receives a data packet it calculates a hash value based on the source and destination information contained in that packet. The calculated hash value becomes the data flow identifier, and all data packets with the same calculated hash value are deemed to belong to the same data flow. If the hash function is not collision free, a sub identifier may be necessary as part of the flow identifier to account for the case where two or more different flows result in the same calculated hash value.

The RMA 116 analyses all packets, either emanating or terminating, and for each packet calculates a flow identifier to determine whether the packet belongs to an existing data flow or a new flow. The flow identifier of an emanating packet is calculated by a hash over the packet's (source address X1 110, destination address Y1 118), and the flow identifier of a terminating packet calculated by a hash over the packet's (destination address X1 110, source address Y1 118). By switching the order of the source and destination address of terminating packets, the hash value of emanating packets and terminating packets being sent between the same network nodes is the same, indicating they are part of the same flow.

The information defining the 'source' and 'destination' addresses of data packets may be decided depending on the level of traffic detail and/or control required. For example, if limited detail and/or control is required the hash values may be calculated on IP addresses only. In this case each flow will be relatively large, denoting all packets being sent from the IP address of X1 110 to the IP address of Y1 118 and all packets from the IP address of Y1 118 to the IP address of X1 110.

Preferably the hash value is calculated on the IP address, the port address and the protocol identifier (e.g. an identifier denoting the file transfer protocol). In this case each flow will be relatively smaller, consisting only of those packets of the same protocol being sent from a particular port on the source IP address of X1 110 to a particular destination port on the destination IP address of Y1 118 and packets from a particular source port on the source IP address of Y1 118 to a particular destination port on the destination IP address of X1 110.

Table 1 sets out a number of exemplary hash value calculation schemes that could be implemented in embodiments of the present invention. Others are also possible.

Address	Emanating flow ID: hash on	Terminating flow ID: hash on	Detail/ Control
IP address	source IP, destination IP	destination IP, source IP	Low
IP address Port address	source IP, destination IP, source port, destination port	destination IP, source IP, destination port, source port	Medium
IP address Port address Protocol ID	source IP, destination IP, source port, destination port, protocol ID	Destination IP, source IP, destination port, source port, protocol ID	High

Table 1

In an alternative but less effective embodiment, flow identifiers of forward and reverse flow components need not be calculated to be the same – i.e. the flow identifier for the emanating packets is calculated by a hash over the packet's (source address, destination address) and the flow identifier for the terminating packets is calculated as a hash over the packet's (source address, destination address). In this way the flow identifier of the emanating packets travelling from X1 110 to Y1 118 will be different to the flow identifier of the terminating packets travelling from Y1 118 to X1 110.

If this is the case the forward and reverse flows may be associated with each other in a list or similar so the RMA recognises they are part of the same flow, or may even be considered and managed as distinct flows by the RMA. If they are managed separately, important information such as the amount of data being sent back into the private network as a result of a particular forward flow is lost. If the forward and reverse flow components are associated with each other at a later stage, e.g. by associating the flows in a secondary list/table, greater computational and memory overhead are introduced.

In a still further embodiment, an estimate of the size of the reverse flow may be made by analysis of the forward flow component e.g., by analysis of the protocol of the forward flow component. For example, if the forward flow data packets are a request for a web page this is likely to require far less traffic in the corresponding reverse flow than if the forward flow data packets are requesting streaming video.

Tokens and token handling

In order to efficiently monitor and manage bandwidth on the available interfaces and make routing decisions, the RMA maintains at lease one (and preferably more than one) token buffer for each of the interfaces. Tokens effectively represent a unit of bandwidth, each token accounting for a fraction of the interface's transmission rate. For example, an interface may have an estimated total transmission capacity of 100 kilobytes per second, and a single token may represent 1 kilobyte per second – in this case the token buffer for the interface would have 100 tokens representing the entire bandwidth capacity of the interface.

Where an interface has dedicated outgoing and incoming bandwidth (i.e. a full duplex connection), forward and reverse token buffers are preferably maintained. If the connection is half duplex, i.e. data may only be sent or received at any given time a single token buffer may be used. The number of token buffers used may also be determined on the basis of how bandwidth allowances are calculated by the ISP (or other entity) to which the interface is connected. If for example, bandwidth limits for incoming and outgoing data are set independently of each other then it is preferable to use a dedicated token buffer for each direction of data

flow. However, if the total bandwidth assigned to the interface is fixed but the relative allocation to forward and reverse flow components can be varied then it may be preferable to use a single shared token buffer to manage bandwidth usage in both directions.

In general terms, a token buffer for an interface has tokens removed from it or added to it to account for fluctuations in the amount of bandwidth being used by the flows being routed through the interface. An entirely unused interface will have a completely full token buffer and an interface for which all available bandwidth has been assigned to one or more flows will have a completely empty token buffer. In use tokens are removed from a token buffer and assigned to flows as they are assigned to the interface or if they increase or decrease in size and tokens are returned to the token buffer if a flow stops (e.g. is timed out) or reduces in size.

As flows are added to and removed from (in the case of the flow timing out) an interface the token buffer associated with that interface is updated accordingly. For example, in a case with dedicated forward and reverse token buffers, if a new flow arrives on a particular interface an initial number of tokens are reserved for each of the forward and reverse flow components. From time to time the size of the forward and reverse flow components will be estimated and, if the flow turns out to be larger than the initial estimate in either direction, further tokens are assigned to that flow component, reducing the number of tokens available for that interface in the direction. Conversely if a flow component is smaller than expected then the number of tokens assigned to a flow component can be reduced. When the flow finishes, all tokens associated with the flow are returned to the token buffer.

In order to monitor the size and continuity of flows or flow components the RMA 116 maintains a flow tracker as described below.

Flow Tracker and Flow Tracking

The RMA 116 maintains flow tracker comprising a hash-based data structure in which flow state information is stored. Figure 2 provides a representation of the data structure 200 of the information fields of the flow tracker. The index of the data structure is the hash value 202 of the flow identifier as discussed above. For

each individual flow the data structure stores the flow identifier 202, source IP address 204, destination IP address 206, source port 208, destination port 210, protocol ID 212, emanating tokens 214, terminating tokens 216, emanating bytes 218, terminating bytes 220, interface ID 222 and time stamp 224.

The time stamp 224 provides information regarding the last time a packet associated with that flow was received at the RMA. A time to live may be set in the RMA, and if the time stamp 224 indicates that the flow is older than that time to live (i.e. that no packets for that flow have been received at the RMA within the selected time) the entry in the flow tracker relating to that flow is deleted. When a data packet is received by the RMA and is associated with an existing flow, the time stamp 224 corresponding to the flow identifier 202 of the packet is updated.

In order to delete flows that are no longer active the flow tracker may order flows according to the time stamp 224. When a packet is received which is part of an existing flow and the time stamp 224 for that flow is updated, the position of that flow in the flow tracker may be moved to the front of the list. This provides for the efficient management of flows in that old flows can simply be deleted from the tail of the list and additional processing is avoided.

The emanating tokens field 214 and terminating tokens field 216 store the number of flow tokens currently assigned to the emanating and terminating flow components respectively. This is discussed in greater detail below in the Token Handling section.

The emanating bytes field 218 and terminating bytes field 220 store information detailing the aggregate number of bytes sent and received in the emanating and terminating components of the flow respectively.

The interface ID field 222 refers to the particular interface on which packets forming part of the flow are routed through.

Figure 3 depicts the process 300 by which the RMA maintains flow information in the flow tracker. The RMA intercepts 302 all data packets being sent from or to the private network (i.e. all emanating and terminating data packets). Each data

packet is detected 304 to be either an emanating data packet or a terminating data packet.

If the packet is determined to be an emanating packet (for example if the source address of the data packet is an address on the private network) the RMA calculates a flow identifier 306 for the packet as:

hash(source IP address, destination IP address, source port, destination port, protocol ID)

If the packet is determined to be a terminating packet (for example if the source address of the data packet is an address outside the private network) the RMA calculates the flow identifier 308 for the packet as:

hash(destination IP address, source IP address, destination port, source port, protocol ID)

In this way emanating and terminating data packets that belong to the same communication flow are associated to the same data flow and same entry in the data structure.

Ideally a non-colliding hash function is used to calculate the flow identifiers, ensuring that each data flow is assigned a unique flow identifier. If, however, the hash function used to calculate flow identifiers allows duplicates (i.e. the hash value calculated for two packets belonging to separate flows may end up the same), collisions have to be resolved in a secondary data structure such as a linked list.

Once the flow identifier for a data packet has been calculated the RMA compares the calculated flow identifier with flow identifiers stored in the flow tracker 310/312 to determine whether the data packet is part of an existing flow or a new flow 314/316. If the calculated flow identifier of the packet occurs in the flow tracker the packet forms part of an existing flow. If the calculated flow identifier of the packet does not occur in the flow tracker the packet is part of a new flow.

New Flow

If the packet belongs to a new flow, a new entry for that flow identifier is created and stored 318 in the flow tracker.

If the packet is an emanating packet the interface ID for that flow is determined 320 according to the interface assignment or routing strategy as discussed below. If the packet is a terminating packet, the interface for that flow is assigned 322 to the interface through which the packet was received.

The RMA then populates the data fields 324 in the flow tracker corresponding to the new flow. The source and destination IP address fields and source and destination port address fields are populated according to the corresponding information in the data packet (again, noting that if the data packet is a terminating packet the source and destination addresses must be switched). The time stamp associated with the flow is also updated according to the time the data packet was received.

The number of tokens assigned to the flow components is determined as discussed below in relation to token handling, and the emanating and terminating token fields are populated.

If the packet is an emanating packet the emanating bytes field is updated according to the size of the data packet (the terminating bytes field left at zero), and if the data packet is a terminating data packet the terminating bytes field is updated according to the size of the data packet (the emanating bytes field left at zero).

Existing flow

If the calculated flow identifier corresponds to a flow identifier existing in the flow tracker, the packet is deemed to be part of an existing flow. In this case the flow ID, source IP, destination IP, source port, destination port and interface ID fields are already known and stored in the flow tracker and do not need to be updated.

The RMA does, however, update the appropriate data fields 326 to maintain up to date information on flow statistics.

If the packet is an emanating packet, the emanating bytes field is updated to be the existing value for that field plus the size of the packet and the terminating bytes field remains unchanged.

If the packet is a terminating packet, the terminating bytes field is updated to be the existing value of that field plus the size of the packet, and the emanating bytes and field remains unchanged.

The time stamp field is also updated to the time the packet was received.

From time to time, and preferably upon receipt of every new data packet the size of the flow component (or flow) is estimated and the number of tokens assigned to the flow component (or flow) from its corresponding interfaces token buffer is recalculated. Upon recalculation of the number of tokens assigned to a flow, the flow tracker data fields 214 and 216 relating to the assigned number of emanating tokens and terminating tokens respectively are updated.

Table 2 below summarises the update actions required for the flow tracker data structure in the event of a packet being received.

Flow tracker field	Packet corresponds to: New emanating flow	New terminating flow	Existing emanating flow	Existing terminating flow
Flow ID	Calculated flow ID	Calculated flow ID	Packet details correspond to existing flow in flow tracker: no update required.	
Source IP	Source IP of packet	Destination IP of packet		
Destination IP	Destination IP of packet	Source IP of packet		
Source port	Source port of packet	Destination port of packet		
Destination Port	Destination port of packet	Source port of packet		
Emanating tokens	Assign as per policy	Assign as per policy	Update	Does not change
Terminating tokens	Assign as per policy	Assign as per policy	Does not change	Update
Emanating bytes	Size of packet	0	Old value + size of packet	Does not change
Terminating bytes	0	Size of packet	Does not change	Old value + size of packet
Interface ID	Selected interface ID	ID of interface through which packet arrived	Already populated as is existing flow	
Time stamp	Time of packet arrival	Time of packet arrival	Time of packet arrival	Time of packet arrival

Table 2

Further manipulation of the data fields in the flow tracker will be discussed below in relation to failover scenarios.

From time to time, and preferably after every packet is received, the RMA updates the token buffer information 328 as discussed above. If the packet is an emanating packet the RMA then routes the packet through the interface associated through the flow the packet is part of 330. If the packet is a terminating packet the RMA routes the packet to the destination node on the private network 332.

Routing Strategies

Routing strategies for emanating data packets (and flows) may be implemented according to the way tokens are assigned to new flows. Forwarding preferences may depend on a number of factors, such as cost, performance, best practice requirements or service types, and strategies may be changed dynamically depending on external factors such as the time of day or traffic thresholds.

Overflow routing

Overflow routing is a strategy that is useful in the case where some interfaces are preferable over others – for example one interface is cheaper than the other interfaces and therefore preferable.

In this scenario one path (for example, the cheapest path) is designated to be the default path and is the first choice for routing new flows on. If that path becomes 'full' – i.e. estimations indicate that no bandwidth is available in either the forward or reverse direction, the new flow is routed to the next preferred interface and so on.

For such a routing scheme when a packet belonging to a new flow arrives the RMA checks the default interface and if sufficient tokens are available for both directions on that interface, it assigns the new flow to that interface (and reduces the tokens in the token buffer(s) accordingly). If, when the default interface is checked, no tokens are available, the next preferred interface is checked for available tokens and, if tokens are available, the flow is assigned to that interface.

Accurate load balancing

If there are no inherent reasons why a particular interface should be preferred over another (e.g. the costs and other overheads associated with all interfaces are the same), the chosen routing strategy may be to distribute traffic evenly between the available interfaces.

This even distribution may be achieved in a number of ways, the simplest of which being when a packet belonging to a new flow arrives the available tokens on each interface are checked and the flow is routed onto the interface having (nominally or proportionally) the most available tokens. Alternatively the new flow can be routed onto the interface which results in the most evenly distributed "interface utilization" across all the possible interfaces. In this case the interface utilization is calculated by:

Tokens used/total possible tokens for interface.

Failover

In the case of one or more interfaces failing, traffic on failed links must be rerouted. The failing of an interface may be detected by the operating system and signalled to the RMA. Where such a signal is received the RMA reduces the number of available tokens for the failed interface(s) to zero and flushes all the flow trackers for flows on that link.

Once the flows are flushed the next packet for that flow arriving at the RMA is not recognised as a packet for an existing flow and is routed as if it is a packet belonging to a new flow.

In this manner only flows that were assigned to the failed interface(s) are impacted, with all other flows remaining on their assigned interfaces.

Although in the preferred embodiment the routing distribution application and all above functionality is described as a single application it is, of course, possible to distribute the functionality between any number of applications and/or physical devices.

Australian Telecommunications Cooperative Research Centre
By Freehills Patent & Trade Mark Attorneys
Registered Patent Attorneys for the applicant
24 May 2006

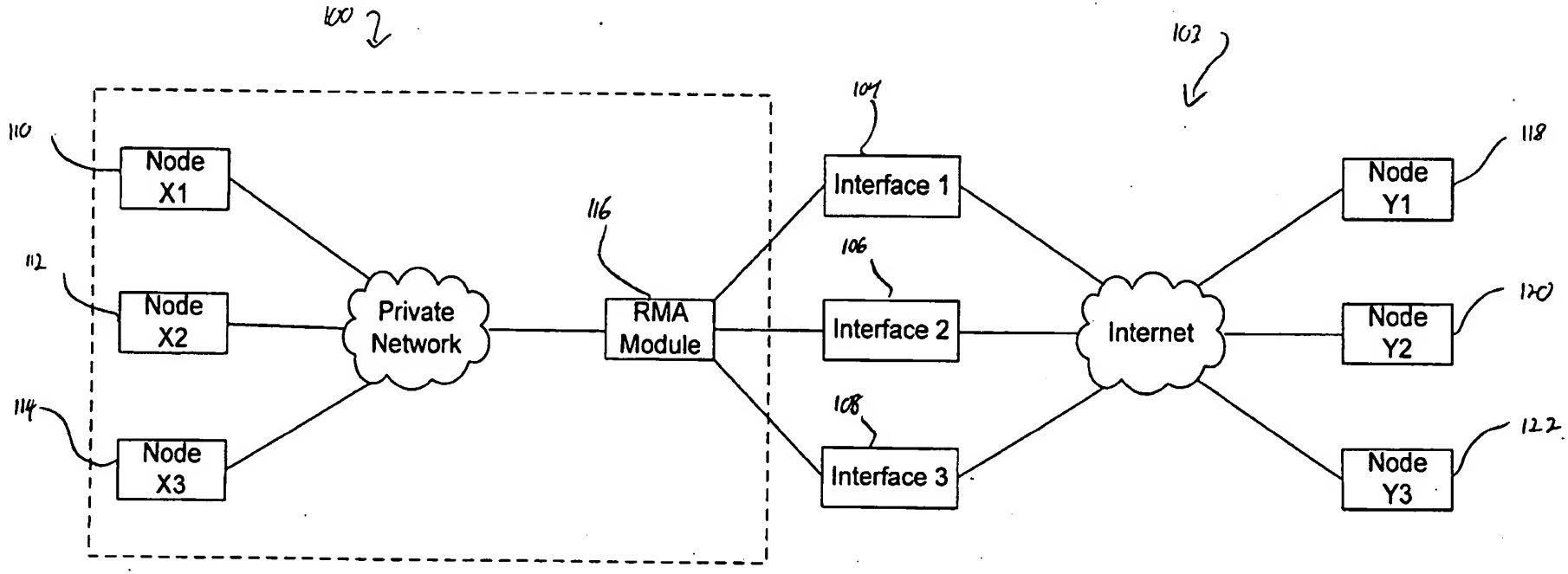
100
102
110
Node X1
112
Node X2
114
Node X3
Private Network
116
RMA Module
104
Interface 1
106
Interface 2
108
Interface 3
Internet
118
Node Y1
120
Node Y2
122
Node Y3

FIG. 1

200

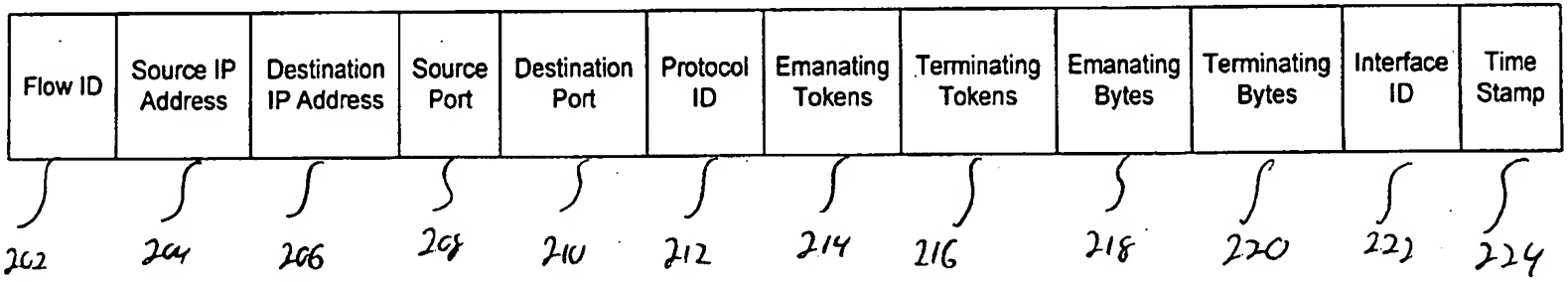
Flow ID
Source IP Address
Destination IP Address
Source Port
Destination Port
Protocol ID
Emanating Tokens
Terminating Tokens
Emanating Bytes
Terminating Bytes
Interface ID
Time Stamp
202
204
206
208
210
212
214
216
218
220
222
224

FIG. 2

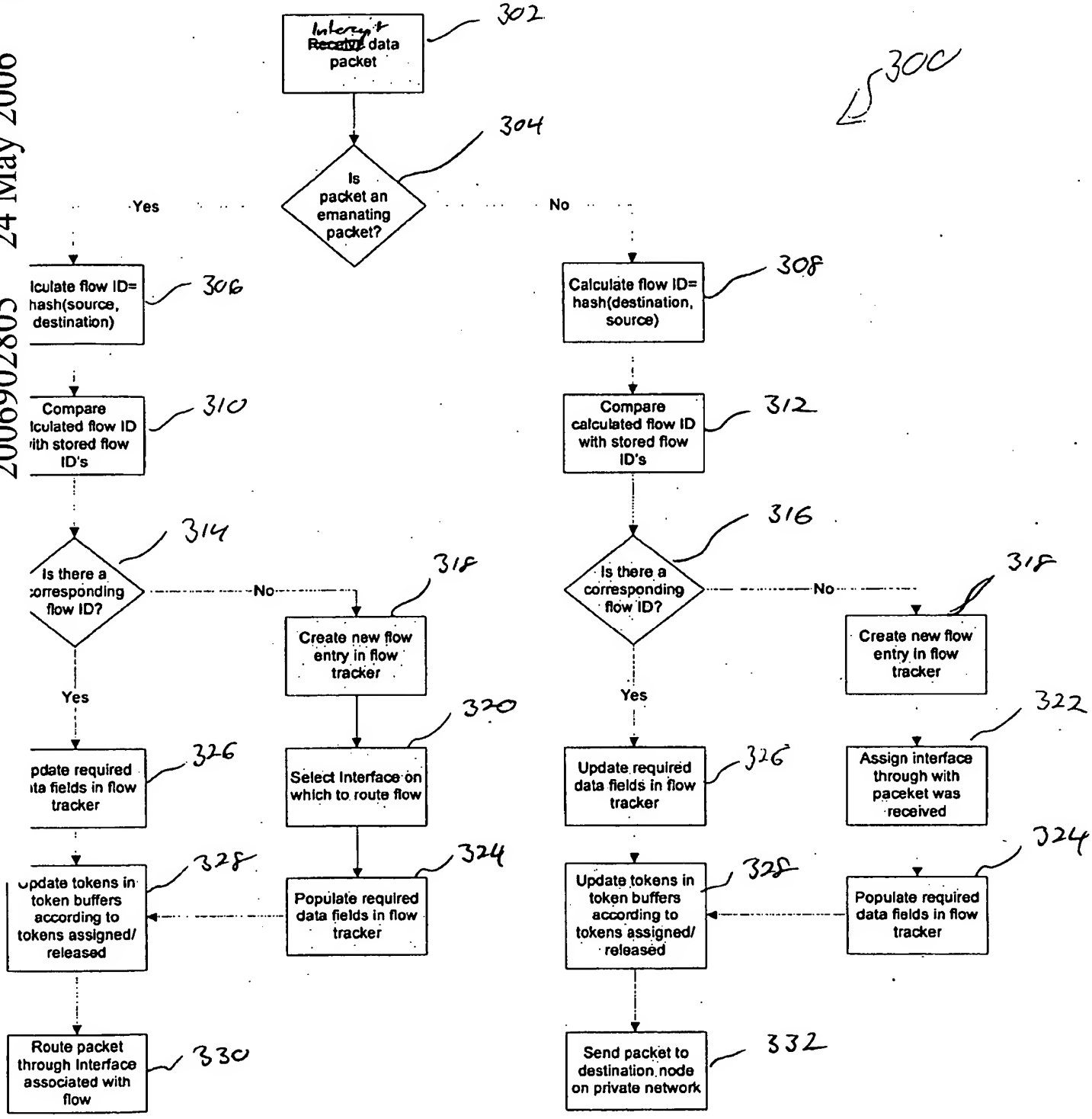

FIG. 3